Cuando el algoritmo discrimina

Raquel Jorge Ricart

¿Cómo se produce la discriminación algorítmica, qué sectores o personas son los más afectados?

“La injusticia, en cualquier parte, es una amenaza a la justicia en todas partes”. Ya lo apuntaba Martin Luther King con los derechos civiles en los 70. Hoy, en 2021, sigue aplicándose y, en ella, entra un factor añadido, novedoso -aunque, mejor dicho, conocido- de los últimos años: el de los algoritmos; o, más bien, la cuestión de la justicia ante la discriminación algorítmica.

Los algoritmos están teniendo un papel cada vez más significativo en asuntos tan esenciales como los servicios públicos: desde la dotación de bonos sociales, la identificación de niveles de riesgo en situaciones de violencia de género, o la automatización de los sistemas de recursos humanos y contratación en empresas, hasta la decisión de cuántas patrullas policiales deben ir a cierto barrio en una ciudad porque se ha detectado que es una zona de mayor peligrosidad.

Por un lado, el sistema algorítmico permite optimizar la toma de decisiones, detectar patrones en situaciones críticas de forma más rápida o procesar cantidades masivas de información para dotar de una solución rápida. Ejemplos claros de ello son el software que identifica disparos en las calles en Estados Unidos, la predicción de aglomeraciones en las estaciones de tren de una ciudad o la entrada por carretera a una zona urbana. Sin embargo, por otro lado, según indica el conocido informe *Poverty Lawgorithms* del Centro Data & Society de la Universidad de Baltimore, la toma de decisiones basada en algoritmos, tal y como se está desplegando ahora, agrega “más alcance, escala y velocidad” (más temas de forma más intensa y plural y de manera más rápida) a las consecuencias negativas de los sistemas tradicionales de distribución y reparto de servicios a la gente en situación de pobreza. A ello se une el hecho de que los

modelos de lenguaje y de procesamiento de datos de los algoritmos incorporan sesgos de género y de raza: varios estudios confirman que el sistema confunde por error a una mujer negra con un hombre negro el 35% de las veces.

Algunas vías de discriminación algorítmica

Un ejemplo claro de ello es la discriminación algorítmica -por errores de datos- en la concesión de hipotecas o la evaluación de inquilinos. En 2019, en el estado de Connecticut (EE UU) una agencia inmobiliaria que disponía de un software de evaluación individualizada de inquilinos para automatizar la concesión de alquileres consideró que una madre no podía alquilar un piso porque su hijo, con diversidad funcional, había sido años antes arrestado por robo erróneamente. El software de la empresa incluía a todas las personas que habían sido arrestadas, pero no si estas habían sido finalmente acusadas o no, o si ese arresto se había retirado por un error, como fue este caso. Pese a la batalla judicial entre una asociación de derechos de la vivienda y la empresa, la Corte del distrito federal determinó que la empresa de evaluación automatizada de inquilinos estaba amparada bajo la regulación de la FHA o Administración Federal de Viviendas. El caso está nuevamente en curso a día de hoy.

No solamente es una cuestión de errores en los datos o de falta de precisión. Incluso con la existencia de datos precisos, la discriminación algorítmica tiene un segundo pilar: el tratamiento diferencial. Cada vez más empresas utilizan tecnologías que automatizan el proceso de selección y contratación de personal. Por ejemplo, se demostró en Amazon que un algoritmo procesaba los datos de contrataciones de los últimos diez años. Así, las nuevas contrataciones eran predominantemente masculinas debido al desequilibrio de género dentro de la industria tecnológica. De forma similar, una empresa estadounidense que ofrece servicios de clase online, aumentaba sus precios por 1,8 en las zonas donde había una mayor densidad de residentes asiáticos.

Pero no todo tiene que ver con los datos en sí. Igual que la inteligencia artificial, la discriminación algorítmica se aplica al procesamiento de lenguaje natural, reconocimiento facial, voz e incluso tonos. Volviendo al proceso de selección, cada vez más se realizan entrevistas de trabajo grabadas en vídeo y automatizadas que no son analizadas por personas, sino que el sistema puntúa al candidato o candidata en base a sus expresiones faciales, tono de voz, uso del lenguaje, nerviosismo o tranquilidad, y otros.

El panorama global de la discriminación algorítmica

Aunque la mayoría de casos registrados y elevados a casos judiciales tiene mayor visibilidad en Estados Unidos (con otros ejemplos como los despidos de profesores en base a una puntuación de desempeño laboral, o el pago de precios de seguros más elevados para vecindarios donde hay minorías de población), lo cierto es que la discriminación algorítmica y su respuesta por parte de la sociedad civil o el sector público ocurre también en otros lugares del mundo.

El informe anual "*Automating Society*" de la organización europea *AlgorithmWatch* pone de manifiesto las tendencias de la toma automatizada de decisiones mediante algoritmos en la Unión Europea y sus Estados miembro. En España, entre otros, el SEPE (el Servicio de Empleo Público de España) utiliza un sistema automatizado para calcular los beneficios por desempleo, así como distribuir las ofertas de empleo, entrevistas y cursos de formación. Si bien se desconoce el impacto de este sistema automatizado y si sus resultados serían distintos a los realizados por una persona, a nivel técnico sí se sabe que existe un reto, que es el de la limitada reconciliación de datos: es decir, en el momento en que haya un error en un dato o una casilla en un archivo, dejará de ser compatible con el resto de archivos y el software lo asignará como nulo. El efecto en el desembolso del paro puede ser importante si no se previene este

riesgo. Asimismo, en Alemania, la ciudad de Mannheim lanzó un proyecto de videovigilancia inteligente que generó controversia porque no hay transparencia algorítmica sobre ante qué movimientos “inusuales” el algoritmo está entrenado para reaccionar y detectarlo como peligroso. En ese sentido, podía dar lugar a falsas alarmas y a una doble discriminación algorítmica, activa y pasiva: la pérdida de tiempo para aquellas personas que sufren emergencias policiales reales y el posible arresto de personas inocentes.

Caso similar en algunos países de América Latina y el Caribe. Si bien la plataforma *Derechos Digitales* afirma que en los países LATAM todavía el grado de desarrollo de las decisiones automatizadas no es muy profundo, sí hay algunos países donde han surgido ejemplos controvertidos. En Venezuela, se empezó a implementar en 2012 hasta hoy un sistema biométrico de seguridad alimentaria que ha afectado de manera específica a las mujeres y a la población LGBTIQ+, concretamente a las personas transgénero. Las leyes vinculadas a derechos civiles y de identidad excluyen a este último grupo de población. Así, al haberse “dataficado” a las personas a la hora de asignar quién podrá comprar alimentos o cuántos en un supermercado u otro, el sistema automáticamente excluye al grupo de población que no reconoce. En ese sentido, una cantidad importante de personas es descartada de forma discrecional -no por el algoritmo, sino por el desarrollador de este.

Mientras que en América Latina y el Caribe, las fases prematuras de discriminación algorítmica empiezan a desplegarse en materia de protección social, como la alimentación, el caso de algunos países de África Subsahariana tiende a relacionarse más bien con los derechos políticos y civiles. Debido a la falta de desarrollo tecnológico, en la mayoría de los casos se buscan productos de empresas procedentes de países extranjeros. Un ejemplo de ello es el sistema de reconocimiento facial implantado por el Gobierno de Zimbabue, a partir de un acuerdo con la empresa china CloudWalk, para su uso por parte de las fuerzas de policía y de seguridad. En un país caracterizado por las protestas sociales, o las restricciones a minorías étnicas o grupos vulnerables, el grado de discrecionalidad a la hora de rendir cuentas sobre daños desproporcionados o injustos es mayor.

De forma similar, en el Sureste Asiático la discriminación algorítmica se aplica en varios sentidos. Sin embargo, destaca el caso del sistema de crédito social de China. Se ha hablado mucho de este sistema, pero lo cierto es que las fuentes apuntan a que no hay uno único, sino que es cada ciudad o región la que implementa su propio marco y aúna datos de sectores distintos (por ejemplo, infracciones en la calle y recibir una amonestación en el trabajo) para dar una puntuación que automatice decisiones en otros (como las posibilidades de entrar en un colegio para un hijo o hija). No existe un patrón único.

Formas de reducir la discriminación algorítmica

Cada vez más organizaciones de la sociedad civil, ONG, los sectores público y privado están trabajando por reducir esta discriminación algorítmica. Entre otras, se plantean varias recomendaciones que podrían ser escalables.

En primer lugar, la toma de decisiones automatizada (o ADM) no debe confundirse con la Inteligencia Artificial. La IA es una tecnología y, como tal, el marco de actuación sobre ella se limita a la calidad de los datos, su precisión, su interoperabilidad, su trazabilidad y explicabilidad y otros. Lo que diferencia al ADM es que, en ella, el margen de respuesta se amplía y entran cuestiones sobre cómo abordar el uso de un sistema automatizado *para un propósito específico, la forma en que se desarrolla, por quién se desarrolla y cómo se despliega*. En ese sentido, la creación de mecanismos de transparencia y de justicia algorítmica es mayor porque requiere que esté presente en cada una de las fases del ADM, tanto en lo que se refiere al sistema algorítmico en sí, como a cualquier tipo de factor ajeno que pueda afectar a este (como el caso del sistema biométrico de seguridad alimentaria y las personas transgénero en Venezuela).

En segundo lugar, hay que actuar sobre los derechos de las personas a la hora de ser afectadas por la discriminación algorítmica. Ya lo apuntaba el propio informe de revisión de los Derechos Humanos del *Facebook Oversight Board*, el Consejo de supervisión y asesoramiento de Facebook formado por personas externas expertas y que incorpora capítulos específicos solo sobre la reparación ante daños desproporcionados (*remedy*) o la necesidad de consentimiento informado. Ahora bien, es importante apuntar que no cabe solo el consentimiento, sino que también debe haber notificaciones cuando se interactúa con un sistema automatizado o cuando un ADM toma una decisión sobre un individuo.

En tercer lugar, se debería centralizar la definición de qué representa un “riesgo” y los niveles existentes. En la Unión Europea, el Libro Blanco de IA de 2020 identificaba dos niveles de riesgo -bajo, alto- en las aplicaciones de IA (que no de las decisiones automatizadas en general). Ante esta propuesta, la Comisión de Ética de Datos de Alemania propuso ampliarlo a cinco niveles de riesgo para dotarle de mayor precisión y “justicia algorítmica”. En la reciente publicación de la Estrategia de IA en abril de 2021, la UE ha ampliado el número de niveles hasta cuatro (riesgo inaceptable, alto, limitado, mínimo). Sin embargo, todavía hoy en día las auditorías de las ADM suelen realizarse de forma limitada y, además, no existe una metodología común de cómo prevenir la discriminación algorítmica.

En definitiva, existe toda una pluralidad de modalidades que producen discriminación algorítmica. Asimismo, los sectores, grupos de población y temas afectados son realmente

diversos. Los mecanismos para abordar esta discriminación son todavía reducidos y, en muchos casos, se limitan a la voluntariedad de las organizaciones a la hora de hacer auditorías o de garantizar que los sistemas de decisión automatizada no producen tales efectos desproporcionados o injustos. Es cierto que cada vez más hay mayor concienciación sobre la discriminación algorítmica. Sin embargo, todavía queda mucho por hacer. No es una cuestión de agendas políticas, sino de agendas de Estado: del Estado del Bienestar, de los Estados democráticos, de la búsqueda de garantías democráticas en aquellos países en proceso de democratización o en búsqueda de hacerlo.

Fecha de creación
28 abril, 2021